

---

# IDENTIFICATION DE VOYELLES EN FONCTION DE LEUR DURÉE

M. WAJSKOP\*

Cet exposé décrit les résultats préliminaires obtenus au cours d'une série d'expériences destinées à mesurer les seuils d'identification de voyelles isolées en fonction de leur durée.

Ce point de départ peut paraître contestable à des linguistes et à des phonéticiens. Les segments sonores isolés ont fait, durant le premier quart de ce siècle, l'objet de recherches intensives mais ont été dédaignées, soit que l'appareillage utilisé fût médiocre, que les résultats fussent mal présentés ou que le montage expérimental négligeât les données élémentaires de la psychologie. De ce fait, la plupart des travaux récents partent d'unités considérées comme naturelles. On peut cependant se demander s'ils n'ont pas envisagé le problème de la perception de la parole à un niveau trop élevé. En effet, comme nous savons peu de choses sur la manière dont le signal acoustique est déchiffré, l'analyse qui s'impose au préalable doit vérifier la présence de tous les éléments redondants contenus dans le signal (forme spectrale, qualité de la voix, indices temporels). C'est dire qu'il faut aborder auparavant ce stade d'une perception que l'on peut qualifier de primaire où les segments sonores seraient appréhendés, ainsi que l'a noté Öhman, suivant un processus de consécution simple.

*Protocole expérimental:* 3 voyelles isolées [i], [a], [u], prononcées par un homme et une femme ont été enregistrées. De ces énoncés, normalisés en intensité et contrôlés auditivement et spectrographiquement, on a réalisé, à l'aide d'une porte électronique, des segments de 4 à 400 ms. Deux bandes-tests de 60 stimuli (3 voyelles x 20 durées) disposés en ordre aléatoire ont été administrées individuellement à 30 sujets. Ceux-ci avaient le choix entre 4 réponses (i — a — u — zéro).

## Expérience I

*I.A:* 10 sujets furent soumis au test de la bande n° 1 (voix d'homme). Des résultats, il ressort un % d'erreurs de 4,07. Cependant, on constate que la totalité des erreurs se concentre au-dessous de 20 ms et que, même à 4 ms, le score de reconnaissance atteint 0.77.

---

\* Université de Bruxelles.

*I.B.*: A 10 autres sujets fut administré le test de la bande n° 2 (voix de femme). Le % d'erreurs fléchit à 2,03. Ici également, les erreurs se concentrent au-dessous de 20 ms et le score à 4 ms atteint 0.80.

*I.C.*: En fonction des résultats obtenus antérieurement, les durées supérieures à 40 ms ont été éliminées et les deux voix ont été mêlées sur une nouvelle bande magnétique comprenant 72 stimuli (2 x 3 voyelles X 12 durées). Les résultats indiquent une montée du % d'erreurs qui atteint 9,86. Le score de reconnaissance à 4 ms tombe à 0.67. Si l'on sépare les 2 catégories de stimuli, Vm et Vf, et que l'on considère que le groupe de 10 sujets a fonctionné différemment selon ces 2 classes, on obtient pour Vm un % d'erreurs de 13,3 (360 observations) et pour Vf, un % d'erreurs de 6,4.

### COMMENTAIRES

- (I) Validité des résultats: Les réponses des sujets sont toujours supérieures à la limite du hasard; la zone d'incertitude est réduite et les réponses atteignent rapidement un score de reconnaissance élevé.
- (II) Score de reconnaissance en fonction de  $F_0$ : La meilleure identification des voyelles prononcées par la voix de femme (pour l'ensemble des expériences, Vm: 9,72% d'erreurs — Vf: 4,72%) semble bien être due au fait que l'on offre à l'auditeur, pour un même laps de temps, un échantillon d'informations plus grand.
- (III) Score de reconnaissance en fonction de la durée: Nous avons étudié la distribution des réponses correctes en fonction des erreurs commises pour chaque condition de durée et pour chaque voyelle.

L'indice d'identification est remarquablement élevé même aux très courtes durées. Exemples:

	a	u	i
4 ms	0.73	0.83	0.63
6 ms	0.75	0.90	0.90
8 ms	0.90	0.88	0.90

Ce résultat doit cependant être considéré avec prudence étant donné l'hétérogénéité des observations ainsi rassemblées et la limitation du vocabulaire soumis aux sujets.

### Expérience II.

Nous nous sommes demandés quelles seraient les réactions d'auditeurs si, placés en face de ces mêmes stimuli, il leur était permis de ventiler leurs réponses sur un choix plus large, en l'occurrence les 7 voyelles du système français minimal (1): [i], [u], [y], [A], [E], [OE] et [O]. Cette expérience ne concerne jusqu'à présent que

10 sujets. Les 3 bandes-tests ont été normalisées aux 12 durées initiales. La matrice globale indique la direction des confusions.

	A	i	u	OE	E	O	y	
A	452	1	3	7	13	4	0	(47%) 480
i	0	422	10	1	30	4	13	(52%) 480
u	7	6	424	2	3	34	4	(61%) 480
	459	429	437	10	46	42	17	440

Fréquence des réponses correctes: 1298 — score d'identification 0.901

Les tendances sont claires: [A] et [i] tendent à être confondus avec [E] tandis que [u] glisse vers [O]. Les voyelles fermées sont donc entendues plus ouvertes, soit moins aigüe, soit moins grave. La voyelle ouverte et centrale est entendue plus fermée et plus antérieure.

Si l'on transcrit graphiquement, (fig. 1) sous forme de résultantes de vecteurs, les tendances ainsi exprimées et ce, à l'intérieur du triangle acoustique décrit par Delattre (2), on observe que les directions des 3 résultantes se croisent dans un espace exigü, proche de la région du schwa. Ceci confirme sur le plan perceptif mais de manière indirecte et non quantitative, les conclusions énoncées par Lindblom (3) sur le plan articulatoire et acoustique.

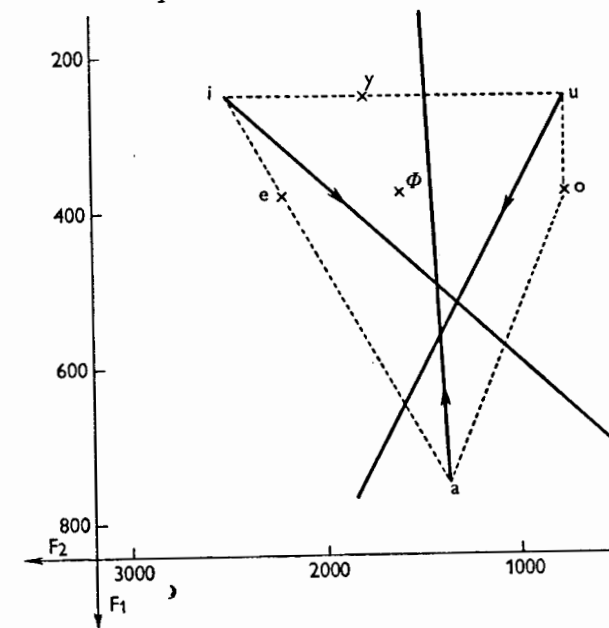


Fig. 1.

## CONCLUSIONS

1. Sur le plan de l'identification, nos résultats confirment ceux déjà obtenus par Gray (4), Peterson (5) et Cramer (6): reconnaissance aisée de vocoides dont la durée n'atteint pas une période complète.
2. (i) La réduction temporelle, à l'intérieur d'un cadre expérimental symétrique, tend à confondre les voyelles les plus fermées (i et u), à perturber la reconnaissance de [a] ou à la confondre indifféremment avec les voyelles fermées d'avant ou d'arrière.  
(ii) L'augmentation de la hauteur tonale améliore les performances des sujets.
3. (i) Par contre, à l'intérieur d'un cadre asymétrique, les voyelles fermées sont entendues plus ouvertes dans leur plan respectif tandis que la voyelle centrale tend à se diriger vers le haut et l'avant.  
(ii) Les scores d'identification pour les courtes durées sont inférieurs à ceux obtenus dans la première série d'expériences.

## REFERENCES

1. B. Malmberg, La structure phonétique de quelques langues romanes. *Orbis*, XI, 1, 131—178 (1962).
2. P. Delattre, Comparing the vocalic features of English, German, Spanish and French. *Iral*, II, 2, 71—97 (1964).
3. B. Linblom, On Vowel Reduction. *Report n° 29, STL, RIT*, Stockholm (1963).
4. G. W. Gray, Phonemic Microtomy: the minimum duration of perceptible speech sounds. *Speech Monographs*, IX, 75—90, (1942).
5. G. Peterson, The Evaluation of Speech Signals. *JSHR*, XIX, 2, 168 (1954).
6. B. Cramer, Über das Erkennen von Sprachlauten. *Aufnahme und Verarbeitung von Nachrichten durch Organismen*, pp. 63—70 (1961).
7. S. Öhman, On the Perception of Swedish Consonants in intervocalic position. *Report n. 25, STL, RIT*, Stockholm (1962).

## DISCUSSION

Rossi:

Le score de reconnaissance très élevé dans votre première expérience est dû à la différence de hauteur des voyelles présentées aux sujets; si cette différence de hauteur n'est pas spécifique de la voyelle, elle doit évidemment être supprimée; mais ne doit-on pas tenir compte de la hauteur spécifique (voir les travaux de Fairbanks et Peterson) qui fait partie intégrante des propriétés des voyelles et présenter par conséquent les voyelles à des hauteurs différentes? Ne risque-t-on pas, dans le cas contraire, de fausser les résultats de l'analyse? Ma remarque vaut également pour l'intensité spécifique.

Wajskop:

Les stimuli ont été normalisés afin de répondre à une demande du département de psychologie expérimentale. Dans les expériences ultérieures, il a été tenu compte des différences spécifiques de hauteur (Peterson-Lehiste) et d'intensité (Fairbanks) afin d'éliminer le caractère "synthétique" de nos stimuli. Dans la Ière expérience, les confusions ne se dirigent pas nécessairement vers les voyelles ouvertes quand il s'agit de [i] et [u]. On note des confusions dans le sens horizontal de [i] vers [u] et l'inverse.